Cortical Responses to Chinese Phonemes in Preschoolers Predict Their Literacy Skills at School Age

Tian Hong, Lan Shuai, Stephen J Frost, Nicole Landi, Kenneth R Pugh & Hua Shu

To cite this article: Tian Hong, Lan Shuai, Stephen J Frost, Nicole Landi, Kenneth R Pugh & Hua Shu (2018) Cortical Responses to Chinese Phonemes in Preschoolers Predict Their Literacy Skills at School Age, Developmental Neuropsychology, 43:4, 356-369, DOI: 10.1080/87565641.2018.1439946

To link to this article: https://doi.org/10.1080/87565641.2018.1439946

Published online: 09 Mar 2018.
Cortical Responses to Chinese Phonemes in Preschoolers Predict Their Literacy Skills at School Age

Tian Hong, Lan Shuai, Stephen J Frost, Nicole Land, Kenneth R Pugh, and Hua Shu

ABSTRACT
We investigated whether preschoolers with poor phonological awareness (PA) skills had impaired cortical basis for detecting speech feature, and whether speech perception influences future literacy outcomes in preschoolers. We recorded ERP responses to speech in 52 Chinese preschoolers. The results showed that the poor PA group processed speech changes differentially compared to control group in mismatch negativity (MMN) and late discriminative negativity (LDN). Furthermore, speech perception in kindergarten could predict literacy outcomes after literacy acquisition. These suggest that impairment in detecting speech features occurs before formal reading instruction, and that speech perception plays an important role in reading development.

Introduction
A large body of research has shown that the ability to recognize, manipulate, and decode basic phonological units is critical for successful reading acquisition (Cao, Bitan, Chou, Burman, & Booth, 2006; Castles & Friedmann, 2014; Hämäläinen et al., 2017; Hulme, Bowyer-Crane, Carroll, Duff, & Snowling, 2012; Ramus & Szenkovits, 2008; Shankweiler & Lundquist, 1992; Ziegler & Goswami, 2005). Although recent research has highlighted the multifactorial nature of developmental dyslexia (DD), with deficits in a range of perceptual and cognitive processes (Pennington, 2006; Pennington, Schlitt, Jackson, Schulz, & Schust, 2012), the core feature that characterizes dyslexia across languages is a deficit in phonological processing (Lyon, Shaywitz, & Shaywitz, 2003; Ramus & Szenkovits, 2008; Snowling, 2001). On this view, phonological deficits that result from suboptimal phonological representations lead to poor decoding (Elbro, 1996; Fowler, Brady, & Shankweiler, 1991; Snowling & Hulme, 1994).

Although the underlying cause of the phonological deficit remains unclear, some research indicates that suboptimal phonological representations in DD arise from deficits in processing speech sound information, including impairments in representation, recognition, and retrieval of speech (Fowler et al., 1991; Muter, Hulme, Snowling, & Taylor, 1998). Indeed, several studies provide support for the phonological deficit in DD arising from lower level, perceptual impairments in speech perception or lower level auditory sensory processing (Boets, Ghesquière, Van Wieringen, & Wouters, 2007; Goswami, 2011, 2015; Steinbrink, Zimmer, Lachmann, Dirichs, & Kammer, 2014; Tallal, 2004). These studies have reported dysfunctions in speech perception and categorization in children with DD (Kraus et al., 1996; Tallal, Miller, & Fitch, 1993), and estimates suggest that at least
one-third of DD children have deficits in perceiving speech sounds and even basic acoustic contrasts (Bishop, 2007; Goswami, 2011; Hämäläinen, Salminen, & Leppänen, 2013; McBride-Chang, 1995; Ramus et al., 2003; Schulte-Körne & Bruder, 2010; Tallal & Gaab, 2006). However, support for deficits in speech perception in DD is mixed, with many studies failing to show differences between dyslexics and controls (Adlard & Hazan, 1998; Démonet, Taylor, & Chaix, 2004; Manis et al., 1997; Ramus et al., 2003; Sperling, Lu, Manis, & Seidenberg, 2003). Indeed, other researchers suggest that rather than deficits in speech perception or auditory processing per se, children with DD may have difficulty in accessing phonemes (Boets et al., 2013) or utilizing phonemic information (Frost et al., 2009; Preston et al., 2016).

In general, children are diagnosed with DD after Grade 2 in elementary school. Research using neuroimaging has shown that children with DD in this age have atypical structure and function in dorsal and ventral components of the reading system (McCandliss & Noble, 2003; Pugh et al., 2001; Sandak, Mencl, Frost, & Pugh, 2004; Schlaggar & McCandliss, 2007). While these studies have been important for identifying the brain bases of DD, research done with school-age children provides a post-diagnostic snapshot rather than a biomarker of risk for DD. Critically, research conducted with younger children, before formal literacy instruction, has revealed neuroanatomical differences in children can be observed in regions known to support phonological processing (Raschle, Zuk, & Gaab, 2012). Further, auditory event-related potentials (ERPs) to speech sounds measured before school-age were also found to be associated with later reading outcomes (Espy, Molfese, Molfese, & Modglin, 2004; Hämäläinen et al., 2017, 2013; Maurer et al., 2009). Following up on this work, we examine relations among neural response to speech (ERPs) and later reading in a cohort of Chinese preschoolers followed longitudinally. Importantly, we include children who are at risk for DD, as indicated by poor phonological awareness (PA) at study entry. The use of PA to define risk is supported by a large literature that has shown that PA predicts successful literacy acquisition and separates DD from typical readers in both alphabetic and nonalphabetic languages (Bryant, MacLean, Bradley, & Crossland, 1990; Ziegler & Goswami, 2005). Moreover, PA can be measured in our sample of pre-readers and will be uncontaminated by reciprocal influences from reading instruction. Although PA has been linked to DD, using preschool PA to define risk, which we term “phonological deficit risk,” rather than family history, is a novel approach, particularly for Chinese.

Much of the work linking early phonological processing ability to reading acquisition or failure has entailed research in alphabetic orthographies. Studying early phonological processing in preliterate children with a nonalphabetic orthography such as Chinese allows associating additional novel properties with both pre-reading phonological skills and later reading skills. One particular property of interest for Chinese is lexical tone. Different from the linguistic usage of pitch variation in nontonal languages, tonal languages like Chinese differentiate lexical meanings by pitch changes. Although lexical tone is a suprasegmental feature, it behaves like segmental features similar to consonants and vowels (Schirmer, Tang, Penney, Gunter, & Chen, 2005). Thus, perception of lexical tone or “lexical tone awareness” may be associated with reading and poorer in Chinese children with dyslexia (e.g., Cheung et al., 2009; Li & Ho, 2011; Shu, Peng, & McBride-Chang, 2008; Zhang et al., 2012). Indeed, Chan and Siegel (2001) reported that Cantonese subjects who had poor reading scores also performed poorly in tone perception. Similarly, Siok and Fletcher (2001) discovered a significant correlation between tone awareness and character recognition. Zhang et al. (2012) found that Chinese children with DD had deficits on categorical perception of lexical tone. These findings suggest that lexical tone perception plays an important role in reading development in Chinese language.

Much of the previous work investigating speech perception in DD has used tasks that require categorizing speech sounds (Blomert, Mitterer, & Paffen, 2004; Cheung et al., 2009; Mody, Studdert-Kennedy, & Brady, 1997; Serniclaes, Sprenger-Charolles, Carré, & Demonet, 2001; Werker & Tees, 1987; Zhang et al., 2012); however, because explicit categorization tasks require attention and motivation, the use of passive electrophysiological designs to measure speech perception in young children is gaining popularity. A number of studies have examined auditory ERPs to speech in
longitudinal studies of reading outcome (Espy et al., 2004; Hämäläinen et al., 2013; Maurer et al., 2009).

In the current study, we focus on two ERPs that have been used to index speech processing and discrimination, the mismatch negativity (MMN) and late discriminative negativity (LDN). The MMN component is an ERP that can reflect discrimination among repeated stimuli and detection of novel stimuli at the pre-attentive stage (Näätänen, 1995; Näätänen, Paavilainen, Rinne, & Alho, 2007) when the processing is automatized even for high-level language processing like syntax and semantics (Pulvermüller & Shtyrov, 2003). In typically developed school-age children, MMN responses to auditory stimuli are similar to adults in latency and topography (e.g., Corbera, Escera, & Artigas, 2006; Hämäläinen, Leppänen, Guttorm, & Lyytinen, 2008; Huttunen, Halonen, Kaartinne, & Lyytinen, 2007; Lachmann, Berti, Kujala, & Schröger, 2005; Sharma et al., 2006). However, the MMN component in younger children has longer latency, compared to adults, despite similar topography (Cheour, Leppänen, & Kraus, 2000). In addition to the MMN, some researchers have observed a second negativity, occurring between 300 and 600 msec after stimulus onset in oddball paradigms, which is named LDN (sometimes also referred to as “late MMN;” Korpilahti, Krause, Holopainen, & Lang, 2001; Kushnerenko et al., 2001). The LDN component is also thought to reflect processing of novel stimuli (Čeponienė et al., 2004; Cheour, Korpilahti, Martynova, & Lang, 2001; Schulte-Körne, Deimel, Bartling, & Rentsch, 2001). Several studies have shown that both MMN and LDN are attenuated in DD (e.g., Corbera et al., 2006; Neuhoff et al., 2012; Schulte-Körne et al., 2001).

In the present study, we measure auditory MMN and LDN in an oddball paradigm as an index of Chinese preschoolers’ speech perception before they began formal schooling (and therefore formal literacy instruction). We divided preschoolers based on PA skills and followed them longitudinally for 1 year to examine both early speech perception and later literacy outcome for those with phonological deficits relative to typically developing children.

We were interested in the following key questions: (1) how does speech perception in preschoolers with phonological deficits compare to those with typical phonological skills? (2) Do phonological deficits affect perception of consonants and lexical tone differently in young children? (3) Can we identify early neural correlates of poor PA that predict later literacy outcome?

Method
Participants

Data were collected as part of a longitudinal study with a cohort of preschoolers, prior to formal literacy instruction, at a mean age of 6.39 years old upon entry to the study (SD 0.31; range 5.60–6.99). A total of 106 Chinese preschoolers were invited to participate in this project. Inclusion criteria were native Chinese speaker, no neurological disease or psychiatric disorders, no ADHD, no impaired sight (uncorrected) or hearing. All children were invited to participate in this longitudinal study before receiving formal instructions (Time 1) and 1 year later (Time 2). Fifty-two of them attended both the ERP experiments and behavioral measurements in Time 1. From the 52 children, we identified a low PA group of 20 children (13 girls, 7 boys)—those with behavioral performance on the PA tasks below the 20th percentile of the performance of the whole longitudinal group (N = 106) on at least two of the four measures of the PA tasks or below the 10th percentile of that on at least one of the four measures. The remaining 32 children (15 girls, 17 boys) we included in a comparison group with the two groups matched on age and nonverbal IQ (based on Raven’s Matrices scores). After the children received formal instruction 1 year later, 43 children (21 girls, 22 boys) received the reading skills tests again. Table 1 presents statistics (mean scores and standard deviations) for the low PA and control group both at entry into the study (Time 1) and after the year of formal literacy instruction (Time 2). No outliers were found in the children who attended the ERP experiments. The distributions for performances in some PA tasks (lexical tone detection, syllable deletion, phoneme deletion) at Time 1 were not normal, so we report the group contrasts (low PA
group vs. controls) using the Mann–Whitney U test for those three tasks in Table 1; the distribution for performances for rime detection, age, and nonverbal IQ were normal, so the group contrasts employed independent samples t-tests. The participants and their parents gave written consent before taking part in the experiments, and they received gift books as compensation. The ethical committee of the Beijing Normal University approved the research protocol.

**Behavioral measures**

**Group identification measures: PA**

Four PA tasks were included to measure Chinese PA. Those tasks were both tested at kindergarten (Time 1) and Grade 1 (Time 2). Performance on those four tasks was the criterion for subgroups (low PA and control group). Mean Z scores of the four tasks (rime detection, lexical tone detection, phoneme deletion, syllable deletion) were adopted as the composite score for PA. Cronbach’s alphas for the mean scores of reading accuracy were $\alpha = 0.66$.

(a) **Rime detection.** The rime detection subtest consisted of 16 trials (Li, Shu, McBride-Chang, Liu, & Peng, 2012). Children listened to one monosyllabic target (e.g., /mao1/ [meaning cat]) and two options (e.g., bao1 [meaning bag] and san1 [meaning mountain]) in one trial. And they were required to choose one from the two options, which sounded more similar to the target. Score was total number correct.

(b) **Lexical tone detection.** The lexical tone detection subtest consisted of 24 trials. Each trial had three monosyllabic words, two of which shared the same lexical tone. Children were asked to pick the one that had a different lexical tone compared to the other two. For example, /chu1/, /shao1/, /wei4/ in one trial, the answer is /wei4/, because the other two share the same lexical tone. Score was total number correct.

(c) **Syllable deletion.** The syllable deletion subtest consisted of 16 trials (Lei et al., 2011). Children were required to produce a new word by deleting the given monosyllable from a disyllabic or trisyllabic phrase. For example, we asked children say /qi4 che1 zhan4/ [meaning bus station] without /zhan4/ [meaning station], and the correct answer was /qi4 che1/ [meaning bus]. Score was total number correct.

(d) **Phoneme deletion.** The phoneme deletion subtest consisted of 26 trials (Pan et al., 2015). Children were asked to say a syllable by deleting the given phoneme from a monosyllabic word. For example, we asked children say /mei4/ without the /m/, the correct answer was /ei4/. Score was total number correct.

### Table 1. Descriptive statistics for behavioral measures and group differences tested at Time 1 and Time 2.

<table>
<thead>
<tr>
<th></th>
<th>Low PA group (N = 20)</th>
<th></th>
<th>Control group (N = 32)</th>
<th></th>
<th>Z/t&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T1 Age (year)</strong></td>
<td>Mean (SD)</td>
<td>Range</td>
<td>Mean (SD)</td>
<td>Range</td>
<td>−0.56</td>
</tr>
<tr>
<td></td>
<td>6.36 (0.32)</td>
<td>5.88–6.99</td>
<td>6.41 (0.32)</td>
<td>5.6–6.93</td>
<td>−0.56</td>
</tr>
<tr>
<td><strong>T1 Nonverbal IQ</strong></td>
<td>23.30 (11.90)</td>
<td>5–43</td>
<td>24.72 (9.73)</td>
<td>12–44</td>
<td>−0.45</td>
</tr>
<tr>
<td><strong>T1 Phonological awareness tests</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>T1 Rime detection</strong></td>
<td>10.25 (3.19)</td>
<td>3–16</td>
<td>12.75 (2.02)</td>
<td>9–16</td>
<td>−3.47**</td>
</tr>
<tr>
<td><strong>T1 Lexical tone detection</strong></td>
<td>7.80 (5.16)</td>
<td>0–20</td>
<td>12.41 (4.66)</td>
<td>6–23</td>
<td>−3.62**</td>
</tr>
<tr>
<td><strong>T1 Syllable deletion</strong></td>
<td>11.50 (3.69)</td>
<td>4–16</td>
<td>14.91 (1.69)</td>
<td>10–16</td>
<td>−3.63**</td>
</tr>
<tr>
<td><strong>T1 Phoneme deletion</strong></td>
<td>1.35 (2.78)</td>
<td>0–9</td>
<td>4.88 (5.04)</td>
<td>0–21</td>
<td>−3.07**</td>
</tr>
<tr>
<td><strong>T2 Character recognition</strong></td>
<td>48.72 (24.03)</td>
<td>12–93</td>
<td>66.72 (31.88)</td>
<td>10–117</td>
<td>−2.1*</td>
</tr>
</tbody>
</table>

Note: *Group contrasts for age, nonverbal IQ, rime detection at Time 1, and character recognition at Time 2 used t-test (t); group contrasts for lexical tone detection, syllable deletion, and phoneme deletion at Time 1 used Mann–Whitney test (Z). *$p < .05$; **$p < .01$.**
**Nonverbal IQ**

*Raven’s Standard Progressive Matrices* (Raven, Court, & Raven, 1996) was used as measure of nonverbal IQ at kindergarten (Time 1). Children were required to choose one fragment that best fits the original picture out of six to eight options. This task contains 5 sets, and each set consisted of 12 items. This task had to be terminated when children’s answers for five consecutive trials were wrong. One point was awarded for each correct answer.

**Literacy skills**

*Character recognition.* This task consisted of 150 Chinese characters, all of which were supposed to have been learned by Grade 6 (Shu, Chen, Anderson, Wu, & Xuan, 2003). Children were required to name all these characters. One point was given for each correct naming.

**ERP stimuli**

Two pairs of Chinese monosyllables were used as stimuli, which were read by a male native Chinese speaker: lexical tone pairs /ji1/ and /ji4/ (Tone 1, the high level tone; and Tone 4, the high-falling tone) and consonant pairs /ba1/ and /ta1/. The lexical tone pairs differed in pitch contour (in fundamental frequency, F0). The acoustic features of F0 in /i1/ were 191.7 Hz and those in /i4/ were onset = 203.6 Hz, end point = 135.6 Hz. The consonant pairs differed in the initial voice onset time (VOT) of consonants. The acoustic features of VOT in /ba1/ were 12.4 msec and those in /ta/ were 42.4 msec. And the first formant (F1) and the second formant (F2) for stimuli /i/ were 340.2/1,907.4 and 808.9 Hz/1,122.3 Hz for /a/. These monosyllables were digitally edited using Sound-Forge (SoundForge9; Sony Corporation, Tokyo, Japan) to have 140 ms duration and 70 dB.

**ERP procedure**

Participants were seated comfortably in an acoustically and electrically shielded room and instructed to ignore the presented sounds while watching a self-selected silent movie. Lexical tone pairs and consonant pairs were presented in four separate experiment sessions. Each session started with 30 standard stimuli, followed by standard stimuli (86.5%, 513 trials) and deviant stimuli (13.5%, 80 trials). Those standards and deviants were pseudo-randomly presented with at least three successive standards between deviants. The stimuli were presented to the participants with a randomly distributed interstimulus interval of 450–500 msec via a loudspeaker located approximately 1 m in front of the participant at 70 dB root mean square intensity level. Lexical tone pairs (/ji1/ and /ji4/) contained two sessions, the first one presented /ji1/ as a standard and /ji4/ as a deviant; the second presented /ji4/ as a standard and /ji1/ as deviant. Consonant pairs (/ba1/ and /ta1/) also contained two sessions in which /ba1/ or /ta1/ was presented as a standard stimulus. The sequence of these four sessions was counterbalanced among participants. The whole experiment lasted for approximately 30 min, and children could have a rest (3–5 min) between each block.

**Electroencephalogram recording**

Continuous EEG was recorded using a Geodesic HydroCel Sensor Net (GSN) consisting of 128 electrodes evenly distributed across the scalp and referenced against the vertex electrode (Electrical Geodesics Inc.) at a sampling rate of 1,000 Hz. EEG signals were amplified by the EGI Net Amp 200 Amplifier with NetStation 4.2 software. The GSN also included the electrodes next to, and below, the eyes for recording horizontal and vertical eye movements. Signals were filtered online with a band
pass filter of 0.05–100 Hz. All the electrodes were physically referenced to Cz (fixed by the EGI system). The impedance of each electrode was kept below 50 kΩ throughout the recording.

**ERP data analysis**

ERP waveform analysis was completed offline using BESA 5.3 (Brain Electrical Source Analysis, Gräfelfing, Germany) and EEGLAB (v13.4.4b) software package (Delorme & Makeig, 2004) running in Matlab R2014b. The raw EEG data were first band-passed at 0.3–30 Hz and then referenced to an average of all electrodes across the scalp. Channels with a flat-line duration higher than 5 sec and those poorly correlated with their interpolated reconstruction based on neighboring channels (correlation thresholds = 0.85) were considered abnormal and rejected. No more than 5.4% (≤7) of the channels were discarded (mean number of rejected channels = 3). The signal was segmented between −100 and 700 msec relative to the stimulus onset and baseline corrected (−100–0 ms relative to stimulus onset). All deviant stimuli (consonant and lexical tone) and only the standard stimuli (consonant and lexical tone) before the deviant stimuli were extracted. Independent component analysis (ICA) was used to identify and remove eye blinks or other movement artifacts. In this study, we adopted Infomax ICA algorithms (Makeig, Bell, Jung, & Sejnowski, 1995) to process EEG data, and we used default parameters as implemented in the *runica* function of the EEGLAB (v13.4.4b) toolbox (Delorme & Makeig, 2004). Artifact rejection for the EEG was performed by a maximum voltage criterion of ±75 V on all scalp electrodes. Analyses were performed on the remaining trials (average non-rejected trials: 115/160 for consonant deviant stimuli, 117/160 for lexical tone deviant stimuli, 116/160 for consonant standard stimuli, 117/160 for lexical tone standard stimuli).

Difference waves (deviant minus standard) were calculated for the MMN and LDN for each deviant. Time windows and electrode pools were selected based on previous literature (Zhang et al., 2012) and visual inspection of the topographies. The mean amplitudes of MMN (250–350 msec after the stimulus) and LDN (500–600 msec after the stimulus) were calculated from the symmetric window of 40 msec around the grand average peak latency. The frontal–central electrodes (F3, F4, FCz, FC3, and FC4) were adopted for statistical analyses. Linear regression analysis was executed with character recognition as a dependent variable, gender and age as control variables, preschool-age PA scores (mean Z scores of PA tasks) and ERPs to lexical tone, and consonant stimuli as independent variables.

**Results**

**ERP experiments at Time 1**

The ERP waveforms from the average signals of frontal–central electrodes (F3, F4, FCz, FC3, and FC4) for the standards, deviants, and the difference waveforms are shown in Figure 1. Mixed measures analyses of variances were run per time window (MMN, LDN) with three factors: group (low PA group, control), phoneme category (consonant, lexical tone), and stimulus type (standard, deviant).

**MMN**

For the MMN, a significant main effect of stimulus type \( (F(1,50) = 48.74, p < .001, \eta^2_p = 0.494) \) was found. The phoneme category × stimulus type interaction was significant \( (F(1,50) = 5.15, p = .028, \eta^2_p = 0.093) \). Importantly, the group × stimulus type interaction was significant \( (F(1,50) = 7.46, p = .009, \eta^2_p = 0.130) \). The main effects of stimulus type were significant both in the low PA group \( (F(1,50) = 7.34, p = .009, \eta^2_p = 0.128) \) and in the control group \( (F(1,50) = 61.31, p < .001, \eta^2_p = 0.551) \) in the subsequent simple effect test. In the follow-up pairwise analysis, we found that the difference
Figure 1. (a) ERP waveforms from the average signals of frontal–central electrodes (F3, F4, FCz, FC3, and FC4) elicited by the lexical tone deviants, the standards, and difference waveforms (deviant–standard) in the control and the low PA groups, respectively; (b) maps display the topographic distribution of the mean amplitudes for the lexical tone contrast in the MMN and LDN analysis windows of the control and the low PA groups, respectively. (c) ERP waveforms from the average signals of frontal–central electrodes (F3, F4, FCz, FC3, and FC4) elicited by the consonant deviants, the standards, and difference waveforms (deviant–standard) in the control and the low PA groups, respectively; (d) maps display the topographic distribution of the mean amplitudes for the consonant contrasts in the MMN and LDN analysis windows of the control and the low PA groups, respectively. (e) MMN and LDN mean amplitude values (vertical bars represent one standard error) from the average signals of frontal–central electrodes (F3, F4, FCz, FC3, and FC4) in each group. Note: *p<0.05; **p<0.01.
between the standard and deviant stimulus in the control group was much larger than that in the low PA group ($t(50) = 2.73, p = .009$). In addition, we conducted planned contrasts on the difference waveforms to directly compare groups on the size of the MMN effect for the lexical tone and consonant contrast conditions (see Figure 1(e)). This analysis revealed a smaller MMN effect for lexical tone in the low PA group compared to that in control group ($t(50) = 2.92, p = .005$), but no group difference for the consonant condition. These results suggest that although MMNs to speech stimuli existed in both groups, the MMN effect was weak in the low PA group compared to that in the control group.

**LDN**

For the LDN, the main effects of phoneme category ($F(1,50) = 14.11, p < .001, \eta^2_p = 0.220$), stimulus type ($F(1,50) = 62.78, p < .001, \eta^2_p = 0.557$), and group ($F(1,50) = 4.33, p = .043, \eta^2_p = 0.080$) were all significant. The phoneme category × stimulus type interaction was significant ($F(1,50) = 5.60, p = .022, \eta^2_p = 0.101$). The group × phoneme category interaction ($F(1,50) = 5.01, p = .030, \eta^2_p = 0.091$) was significant. And in the subsequent simple effect test, in the low PA group, the main effect of phoneme category was not significant ($F(1,50) = 0.94, p = .338, \eta^2_p = 0.018$), while in the control group, the main effects of phoneme category were significant ($F(1,50) = 23.35, p < .001, \eta^2_p = 0.318$). It suggested that lexical tone and consonant contrasts induced different auditory potentials in control group, yet stimulus type did not interact with these two factors. Hence, we do not discuss phoneme category differences in the control group further. More importantly, the group × stimulus type interaction ($F(1,50) = 5.35, p = .025, \eta^2_p = 0.097$) was significant. In the subsequent simple effect test, both in the low PA group ($F(1,50) = 12.79, p = .001, \eta^2_p = 0.204$) and the control group ($F(1,50) = 68.10, p < .001, \eta^2_p = 0.577$), the main effects of stimulus type were significant. In the follow-up pairwise analysis, we found that the difference between the standard and deviant stimulus in the control group was much larger than that in the low PA group ($t(50) = 2.31, p = .025$). In addition, we also conducted planned contrasts on the difference waveforms to directly compare groups on the size of the LDN effect for the lexical tone and consonant contrast conditions (see Figure 1(e)). This analysis revealed a smaller LDN effect for consonant stimuli in the low PA group compared to that in the control group ($t(50) = 2.10, p = .040$), but no group differences for the lexical tone condition.

**Prediction to literacy skills at Time 2**

To assess contributions of PA and auditory processing to literacy skill measured at Time 2, hierarchical regression analyses were executed with character recognition at Time 2 as a dependent variable, all independent variables were forced (entered) into the model. In this model, gender and age as control variables were entered in step 1, preschool-age PA scores (mean $Z$ scores of PA tasks) were entered in step 2, and ERPs to lexical tone and consonant stimuli were entered in step 3 (see Table 2). The control variables had a 14.8% contribution, and PA could explain a larger percentage ($\Delta R^2 = 16.3%$). Moreover, the auditory processing at Time 1 also significantly predicted to literacy skill at Time 2 ($\Delta R^2 = 15.9%$). In other words, PA and speech perception at kindergarten predicted future literacy outcomes.

**Discussion**

In this study, we examined speech perception of the low PA preschoolers and controls, and its influence on their future reading skills. We found reduced difference waveforms for the MMN (induced by lexical tone) and the LDN (induced by consonant) for both consonant and lexical tone stimuli in the low PA group compared to the control group, suggesting a low-level deficit in speech
perception of preschoolers. Specifically, this study found that the neural-level indices of risk in preschool, defined here by poor PA, could predict literacy outcomes in later grade. Overall, our results are consistent with what has been observed in previous studies conducted in alphabetic languages and support the hypothesis that speech perception plays a critical role in reading development (Preston et al., 2016; Pugh et al., 2013). Given that phonological processing both predicts and is influenced by learning to read, it is especially noteworthy that we see this pattern linking phonological abilities and speech in pre-readers.

We found significant differences between the low PA and control groups in different stimuli (lexical tone and consonant). Given that reduced MMN and LDN to speech reflect poor discrimination and detection of speech stimuli (e.g., Cheour et al., 2001; Maurer, Bucher, Brem, & Brandeis, 2003), the current results suggest that children who have a deficit in PA also have impairments in speech perception and that these deficits are present before formal reading instruction. It suggests that the interaction between PA and speech perception happens at a very early stage. Preschoolers (Boets et al., 2010; Boets, Wouters, Van Wieringen, & Ghesquiere, 2007; Espy et al., 2004; Gerrits & De Bree, 2009; Hämäläinen et al., 2017, 2013; Law, Vandermosten, Ghesquiere, & Wouters, 2017; Lovio, Näätäinen, & Kujala, 2010; Maurer et al., 2009; Plakas, Van Zuijen, Van Leeuwen, Thomson, & Van Der Leij, 2013) and even infants (Guttorm, Leppänen, Hämäläinen, Eklund, & Lyytinen, 2010; Guttorm et al., 2005; Leppänen, Pihko, Eklund, & Lyytinen, 1999; Pihko et al., 1999; Richardson, Leppänen, Leivo, & Lyytinen, 2003) who have risk for DD are found to have impairments in speech perception. This result is in line with another recent longitudinal investigation of pre-reading English-speaking children with temporal auditory processing measurements (Law et al., 2017). Law and colleagues found significant differences between DD and non-impaired readers on measures of PA across 3 years and that auditory processing in kindergarten predicted later literacy skills.

Previous studies have suggested that speech perception modulates development of reading skills (Goswami, 2011; Goswami et al., 2002; Guttorm et al., 2010; Law et al., 2017; Molfeese, 2000; Thomson, Fryer, Maltby, & Goswami, 2006). For example, Molfeese (2000) found that the auditory ERP responses to speech and nonspeech stimuli in newborns can predict their reading problems after 8 years. Guttorm et al. (2010) also found that newborn responses to syllables are related to phonological abilities, verbal short memory, and reading abilities in preliterate children. The present study replicates and extends these previous studies, which are mostly carried out in alphabetic languages, by examining both suprasegmental and segmental features and revealing the nature of brain–behavior relations. That is, the amplitude of the difference waveforms for MMN and LDN, which were induced by lexical tone and consonant stimuli, was significantly correlated with later measures of literacy skills (character recognition, see in Appendix A). Thus, individuals with better neural responses to speech stimuli before formal instruction have higher reading proficiency 1 year.

### Table 2. Hierarchical multiple regression predicting character recognition at Time 2 (Significant variables are marked in bold).

<table>
<thead>
<tr>
<th>Variable</th>
<th>$B$</th>
<th>$SE$</th>
<th>Beta</th>
<th>Significant</th>
<th>Adjusted $R^2$</th>
<th>$\Delta R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.085</td>
<td>0.148</td>
</tr>
<tr>
<td>Gender</td>
<td>9.97</td>
<td>8.67</td>
<td>0.17</td>
<td>0.257</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 1</td>
<td>4.16</td>
<td>69.85</td>
<td>0.04</td>
<td>0.953</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 2</td>
<td>31.86</td>
<td>67.92</td>
<td>0.32</td>
<td>0.642</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.241</td>
<td>0.163</td>
</tr>
<tr>
<td>T1_PA</td>
<td>18.49</td>
<td>6.08</td>
<td>0.41</td>
<td>0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.350</td>
<td>0.159</td>
</tr>
<tr>
<td>T1_Consonant MMN</td>
<td>5.75</td>
<td>3.20</td>
<td>0.28</td>
<td>0.081</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1_Consonant LDN</td>
<td>1.42</td>
<td>3.23</td>
<td>0.07</td>
<td>0.662</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1_Tone MMN</td>
<td>−0.69</td>
<td>3.83</td>
<td>−0.03</td>
<td>0.859</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1_Tone LDN</td>
<td>−5.97</td>
<td>2.71</td>
<td>−0.33</td>
<td>0.034</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
later. These findings support the explanations for the associations between speech sound processing in young children and future’s reading skill (e.g., Hulme et al., 2012; Ziegler & Goswami, 2005).

The regression analyses of preliterate ERP responses including MMN and LDN components of lexical tone and consonant demonstrated that speech perception in kindergarten speech perception ability also predicts growth in measures of literacy skill (Chinese character recognition). The lexical tone LDN predicted literacy skill when kids were in Grade 1, even after controlling for the relevant behavioral performance (PA) in kindergarten. This result demonstrates a link between early perception in lexical tone and development of Chinese character recognition. Moreover, the current study compared the lexical tone contrasts with consonant contrasts to address a question surrounding the specific role of lexical tone. However, we did not observe a main effect of phoneme category and group differences were similar for both in the lexical tone and consonant conditions, revealing that lexical tone perception and consonant perception have a similar relationship to PA. This result is consistent with other studies investigating the influence of both suprasegmental and segmental features on semantic processing (Schirmer et al., 2005).

**Conclusion**

The current longitudinal study goes beyond previous studies by demonstrating that ERP indices of speech perception in Chinese preschool children can predict reading skill at school age. We find that Chinese preschoolers with low PA have atypical speech perception at both the suprasegmental and segmental levels compared to controls. Moreover, the neural-level indices of risk factors before formal instruction were associate and predictive of later linguistic skills. This work provides empirical evidence for the associations between low-level speech sound processing and reading development.

**Acknowledgements**

This research was supported by National Key Basic Research Program of China (2014CB846103), by Natural Science Foundation of China (31671126, 31611130107), by Beijing Municipal Science & Technology Commission (Z151100003915122), by the Fundamental Research Funds for the Central Universities (2017XTCX04), by the Interdiscipline Research Funds of Beijing Normal University, and by the National Institute of Child Health and Human Development (grant number P01 HD001994) to Haskins Laboratories.

**References**


Hämäläinen, J., Landi, N., Loborg, O., Lohvansuu, K., Pugh, K., & Leppänen, P. H. T. Brain event-related potentials to phoneme contrasts and their correlation to reading skills in school-age children. *International Journal of Behavioral Development, 0*(0), 0165025417728582. doi:10.1177/0165025417728582


Appendix A. Supplementary material

As can be seen in Table A.1, the consonant LDN showed a significant correlation with PA ($r = -0.332, p = .030$). The lexical tone MMN also showed a correlation with PA ($r = -0.412, p = .006$). The tone LDN showed significant correlations with literacy skills, character recognition ($r = -0.329, p = .031$).

Table A.1. Correlations between ERP amplitudes and behavior in Time 2 (Significant variables are marked in bold).

<table>
<thead>
<tr>
<th></th>
<th>Consonant MMN</th>
<th>Consonant LDN</th>
<th>Tone MMN</th>
<th>Tone LDN</th>
</tr>
</thead>
<tbody>
<tr>
<td>T2_Phonological awareness</td>
<td>$-0.136$</td>
<td>$-0.332^*$</td>
<td>$-0.412^{**}$</td>
<td>$-0.157$</td>
</tr>
<tr>
<td>T2_Character recognition</td>
<td>$0.176$</td>
<td>$-0.077$</td>
<td>$-0.228$</td>
<td>$-0.329^*$</td>
</tr>
</tbody>
</table>

Note: *$p < .05$; **$p < .01$. 